Université de Genève

Département d'informatique

Théorie de l'information pour la science des données et l'apprentissage automatique

Série d'exercices 6 — 03.05.2023 Délai : 09.05.2023

Note : Veuillez suivre les instructions suivantes
`https://moodle.unige.ch/mod/page/view.php?id=171850`.

Entropie relative

L'entropie relative ou divergence de Kullback-Leibler (KLD) entre des PMF $p(x)$ et $q(x)$ est définie (en bits) par:

$$D_{KL}(p||q) = D(p||q) = \sum_{x \in \mathcal{X}} p(x) \log_2 \frac{p(x)}{q(x)} \tag{1}$$

$$= E_{p(x)} \left[\log_2 \frac{p(x)}{q(x)} \right] \tag{2}$$

Propriétés de l'entropie relative *(Prouvées plus bas)*

- $D(p||q) \geq 0$
- $D(p||q) \neq D(q||p)$ en général

Problème 1 (40 points).

Soient X et Y deux variables aléatoires binaires et $Z = X \oplus Y$, où $\oplus$ dénote l'opération XOR.

X	Y	$X \oplus Y$
0	0	0
0	1	1
1	0	1
1	1	0

1. Montrez que $H(Z|X) = H(Y|X)$.
2. Sous quelles conditions $H(X,Y) = H(X) + H(Y)$?
3. Montrez que si X et Y sont indépendants, $H(Z) \geq H(X)$ et $H(Z) \geq H(Y)$.

Problème 2 (30 points).

Soit X une variable aléatoire avec trois valeurs possibles : $\mathcal{X} = \{a, b, c\}$. Imaginons deux distributions pour X: $p(x)$ et $q(x)$.

x	$p(x)$	$q(x)$
a	1/2	1/3
b	1/4	1/3
c	1/4	1/3

1. Calculez $H(p)$ et $H(q)$.
2. Calculez $D(p||q)$ et $D(q||p)$.
3. Ces deux valeurs ne sont pas égales. Sous quelle condition $D(p||q) = D(q||p)$?

Problème 3 (30 points).

1. Soient deux ensembles de nombres non-négatifs $a_1, .., a_N$ et $b_1, .., b_N$. Prouvez que :

$$\sum_{i=1}^{N} a_i \log\left(\frac{a_i}{b_i}\right) \geq \left(\sum_{i=1}^{N} a_i\right) \log \frac{\sum_{i=1}^{N} a_i}{\sum_{i=1}^{N} b_i}$$

   Indice : Vous pouvez utiliser l'inégalité de Jensen

$$\sum_i \lambda_i f(t_i) \geq f\left(\sum_i \lambda_i t_i\right) \qquad \text{où} \sum_i \lambda_i = 1$$

   pour la fonction convexe $f(x) = x \log(x)$.

2. Supposons que $p(x)$ et $q(x)$ soient deux distributions pour X. Prouvez que :

$$D(p||q) \geq 0$$

   Discutez la condition d'égalité.

   Indice : L'inégalité de Jensen est saturée (égalité) si et seulement si t_i est constant.